Anyway, let's have a look at some of the summary data given. Collectively, for wives of non-smokers we have 32 lung cancers among 21 895 women versus 86/44 184 for ex-smokers or light smokers versus 56/25 461 for heavy smokers. That would yield a single df χ^2 by the method Dr Hirayama apparently used of 3·31, or a χ value of 1·82, not significant, two tailed, at the 5% level. Dr Hirayama, however, reports a χ value of 3·30, with a two-tailed probability level beyond 0·1%. The question then is whether he has conducted a more refined analysis, about which he is giving us no clues, or he has mistakenly interpreted his χ^2 value as a χ value.

As an example of a particularly striking association, Dr Hirayama concentrates on the results for non-smoking wives of agricultural workers aged 40-59. The corresponding data for this group are 3/5 999 versus 20/12 753 versus 16/7 150, which would yield a χ^2 of 6·45, a χ value of 2·54. Dr Hirayama reports a surprisingly close χ value of 2·60, with an associated significance level of 0·94%. As an extreme outcome of several possible analyses available to him this is not really all that significant. But if the difference between 2·54 and 2·60 reflects only rounding errors in calculations rather than distinctly different analyses, this would make it more likely that Dr Hirayama's 3·30 value is a χ^2 rather than a χ, which, as such, does not have the extreme statistical significance which he has attached to it.

Much more careful analysis of the data would be needed before it can be claimed that a passive effect of smoking has been clearly established.

NATHAN MANTEL

Department of Statistics,
Biostatistics Center,
George Washington University,
Bethesda, Maryland 20014,
USA

To Dr Takeshi Hirayama

In a recent press release the Tobacco Institute in the United States challenged the statistical validity of your finding that non-smoking wives of heavy smokers have a higher risk of lung cancer. The central point raised by the Tobacco Institute was that you erred fatally in your calculation of a critical test statistic and therefore that your claim of a high level of statistical significance is wrong. The Tobacco Institute's assertions relied heavily on a speculation contained in a memorandum on your study by Dr Nathan Mantel. The purpose of this letter is to inform you of the details of Dr Mantel's analysis, and in particular the origin of the alleged arithmetical error. Your results, at least those published, do indeed display a high level of statistical significance. If there is an error of inference in your study, it is far less superficial than the alleged arithmetical mistake publicised by the Tobacco Institute.

The essential data are given in the attached table, which we derived from tables I and II of your paper in the *B.M.J.* (17 January, p 183). The population of non-smoking wives at the beginning of the prospective study and the number of lung cancer deaths during 14 years of observation are broken down according to the age, occupation, and smoking habits of their husbands. Also provided are the corresponding 14-year lung cancer death rates for each cell. At the bottom of the table the data for all age and occupational categories of husbands are combined.

In his memorandum to the Tobacco Institute, Dr Mantel calculated a χ^2 test statistic with one degree of freedom from the *combined* data, and obtained the result $\chi^2 = 3{\cdot}31$, which would yield a two-sided $p = 0{\cdot}07$. In performing this test he used the scoring 0, 1, 2 for the three smoking categories. This value of χ^2 looked suspiciously close to the value of $\chi = 3{\cdot}299$ reported on page 183 of your paper. On this basis, Dr Mantel *speculated* that you may have failed to take a simple square root in your calculation.

If, however, a χ^2 test statistic with one degree of freedom is calculated from the *disaggregated* data in the table, then the result is $\chi^2 = 8{\cdot}09$, which yields a highly significant two-sided $p = 0{\cdot}004$. In making this calculation we used formulas (2) through (4) on page 694 of Dr Mantel's original paper,[1] which shows how to combine χ^2 statistics from the four independent contingency tables. The reason for the increase in statistical significance is that the dose-response relation between the husband's smoking habit and the non-smoking wife's lung cancer rate, which is apparent from the combined data, becomes even more prominent when the data are broken down by age and occupational categories of the husband. That is exactly what Dr Mantel's extension χ^2 was designed to test.

We reach the same conclusion of a high level of statistical significance when the disaggregated data are fitted to a number of plausible statistical models. For example, if P_{ij} denotes the 14-year lung cancer death rate of a non-smoking woman whose husband belongs to age-occupational category $i = 1, 2, 3, 4$ and smoking category $j = 1, 2, 3$, then the traditional relative risk model $P_{ij} = a_i \cdot r_j$ yields a good statistical fit with an equally strong level of significance. (The 6-degree-of-freedom χ^2 goodness-of-fit statistic is 5·3, with $p = 0{\cdot}5$.) For this model, the estimated relative risk of lung cancer for the intermediate category of smoking husbands is 1·42 (95% confidence interval 0·95 to 2·13) by maximum-likelihood methods. The estimated relative risk of lung cancer for the heavy-smoking husbands is 1·87 (95% confidence interval 1·21 to 2·91). These estimates are consistent with the relative risks reported in your paper.

In fairness and respect to Dr Mantel, we wish to point out that his speculation of an arithmetical error was only a minor point in his memorandum. The greater part of Dr Mantel's criticisms was devoted to ambiguities in your method of presenting the results. Thus it is not entirely clear to us whether you based your statistical calculations on more finely disaggregated age and occupational groups than those reported in the paper and reproduced by us here. It is unclear whether you used the number of person-years at risk or the number of persons at risk in your calculations of statistical significance. It is also not clear to us whether you standardised on the ages of the wives themselves. Such calculations, if not already performed, would certainly make the analysis more conclusive. But we see no obvious reason why these additional analyses should produce substantially different results.

The Tobacco Institute's claim that your study is invalidated by a trivial arithmetical error has become an international news story. We recognise that you, as author of the paper under challenge, are in the most appropriate position to respond publicly. As your colleagues in the scientific community, we urge you to uphold the validity of your paper.

JEFFREY E HARRIS
WILLIAM H DUMOUCHEL

Department of Economics,
Massachusetts Institute of Technology,
Cambridge, Mass 02139,
USA

[1] Mantel N, Haenszel W. *Journal of the American Statistical Association* 1963;58:611-27.

SIR,—The far-reaching implications of the paper on passive smoking (17 January, p 183) make it imperative to determine whether the data on which it is based are adequate for the purpose.

In none of his several reports on this population does Hirayama explain how he selected his study population. There are 46 prefectures in Japan. Within these prefectures there are 832 health stations. In the six prefectures from which the 69 health stations were selected there are 173 health stations. How were the 69 selected? In one report it was stated that they were the medically adequate ones, which would introduce certain biases. The six prefectures included in the study contain most of the heavy industries in Japan and also border the sea. Shipbuilding, for example, in which Japan leads in world production, is associated in the United States with asbestos exposure. Since shipbuilding is a major industry in one of the prefectures in the study, were the health stations selected for the study for residents in the shipbuilding areas? Industry in Japan supplies medical care to workers and their families as one of the perquisites of employment. Were the families covered by industry omitted from the study? If so, the healthiest segment of the population was omitted.

The figure of 91·99%, of the population of the health stations is given to demonstrate the population basis of these data. To what population is Hirayama referring? Is the census in Japan taken of health stations—that is, of individuals eligible for care in health stations? Is not attendance at the health stations the option of the individual? In the *Japanese Almanac* for 1972, the proportion of females visiting the health stations ranges from 2·5% in Miyagi to 14·5% in Osaka. How can this be population based? And if it is population based why were the rates figured by the person-year method rather than on the actual population?

Personal interview surveys usually elicit more complete response than mailed questionnaires or hospital record searches. In this study, on the pivotal question of occupation 44 357 (31·04%)

*Lung cancer deaths in non-smoking Japanese wives during 14 years of observation: data based on tables I and II of Hirayama's paper**

Husband's smoking habit:	Non-smoker	Ex-smoker or 1-19 cigarettes/day	≥20 cigarettes/day
Disaggregated data			
Husband aged 40-59 years and working in agriculture:			
Population of wives at risk	5999	12 753	7150
No of deaths from lung cancer	3	20	16
Lung cancer deaths per 10 000	5·0	15·7	22·4
Husband aged 40-59 years and working elsewhere:			
Population of wives at risk	8021	17 923	13 434
No of deaths from lung cancer	8	20	20
Lung cancer deaths per 10 000	10·0	11·2	14·9
Husband aged 60 years or more and working in agriculture:			
Population of wives at risk	4407	7291	2241
No of deaths from lung cancer	14	32	8
Lung cancer deaths per 10 000	31·8	43·9	35·7
Husband aged 60 years or more and working elsewhere:			
Population of wives at risk	3468	6217	2636
No of deaths from lung cancer	7	14	12
Lung cancer deaths per 10 000	20·2	22·5	45·5
Combined data			
Husbands of all ages and occupations:			
Population of wives at risk	21 895	44 184	25 461
No of deaths from lung cancer	32	86	56
Lung cancer deaths per 10 000	14·6	19·5	22·0

* *BMJ* 1981;282:183-5.

2023381522

Source: https://www.industrydocuments.ucsf.edu/docs/lznx0000

women surveyed and 6245 (40·19%) of the women smokers were listed as unknown or not specified as to occupation. Was this study part of a larger study collected for a different purpose, a national health survey perhaps? This could explain this large number of unknowns in this important category.

It is unusual to have such a study based on the ages of the the husband at time of interview rather than on age of the study subjects, the wives. When the agricultural and non-agricultural comparisons were made, the experience of the wives of men over 69 was omitted, an important segment of the study population. Why was this? Clarification of these points would be appreciated.

A recent study in Japan by Minowa *et al*[1] considers the distribution of lung cancer mortality and many of the environmental factors in Japan. In this carefully conducted and explained study, the author states, "Smoking did not seem to explain the prefectural differences of lung cancer mortality." There was no significant difference between the prefectures with high and low consumption of tobacco. The high standardised mortality ratios were along the seacoast, in the highly polluted and heavy industry areas for both males and females. The areas in Hirayama's study were nearly all in the areas with the high standardised mortality ratios in the Minowa study. Aoki,[2] another reputable epidemiologist from Japan, reported in 1980 on the incidence of lung cancer and its mechanism as regards factors other than smoking. Before the second world war lung cancer was a rare disease but during 1947–60 the number of deaths increased rapidly. Such an increase cannot be explained by smoking. These authors make the point that Japan has a disproportionate amount of adenocarcinoma, and Hirayama also mentions the large number of adenocarcinomas in the cases from this study that had a postmortem examination. The proportion of adenocarcinomas of the lung in Japan, twice as high as in any of the developed countries in the world and validated by an outstanding Japanese thoracic specialist, Ishikawa,[3] indicates aetiological factors peculiar to Japan, specifically related to adenocarcinoma.

Since histology is critical in determining prognosis, treatment, and aetiology in lung cancer, and adenocarcinoma is a different disease from squamous-cell carcinoma, and is omitted from many studies on results of smoking because the relationship is not established, to attribute causal dependence to the single aetiological factor of passive exposure to tobacco smoke regardless of histology raises serious questions.

ELEANOR J MACDONALD

University of Texas System Cancer Center, M D Anderson Hospital and Tumor Institute, Houston, Texas 77030, USA

[1] Minowa M, Shigematsu I, Nagai M, Fukumoto K. *Soc Sci Med* 1981;15D(1):225–31.
[2] Aoki K, Ohno Y. *Nippon Rinsho* 1980;38:2541–50.
[3] Ishikawa S. *Jpn J Clin Oncol* 1973;3:19–30.

SIR,—The validity of the significance test used in my paper was kindly confirmed by prominent statisticians in many institutes, including the US National Cancer Institute and the Massachusetts Institute of Technology.

Professor Macdonald apparently misunderstood the method used in our study. A complete interview survey was done for 265 118 adults aged 40 and above at the time of (or in some cases just subsequent to) the national census in 1965 in 29 health centre districts in six prefectures. This was not a random survey but the satisfactory representativeness of the study samples with regard to demographic and social indices was confirmed after the survey. Asbestos exposure is quite unlikely to have influenced our result since the proportion of industrial areas are surely not in excess in our sample. The husbands' ages are closely associated with wives' ages. Since smoking habit is a characteristic of the husbands, the husband's age was used. We have tables by wives' ages separately.

The age-occupation-specific mortality asked for by Mr Kornegay and Dr Kastenbaum is shown in the accompanying table I. The statistical test on the overall results revealed a highly significant association.

Garfinkel's results[1] in the United States, which failed to show a significant dose-response relationship between amount of cigarette smoking by the husbands and the mortality of their non-smoking wives from lung cancer, are not a surprise to me. The discrepancy could be partly due to different methodology. Our study in Japan is based on census population and husbands as well as wives were interviewed with a 95% response rate. Since Garfinkel's series was based on volunteers' interviews, the husband's smoking history was available only for 27·1% of the non-smoking wives with lung cancer, 153 out of 564, in sharp contrast to 72·3% in our series (174 out of 240).

When, however, one compares the three recent studies—those of Garfinkel,[1] Trichopoulos *et al*[2] and ours—the difference between the actual and the expected mortality from lung cancer in non-smoking wives of smoking husbands were all in the same direction (table II); and a sound researcher will conclude that the size of the study samples should be increased, especially in the US series.

Other possible reasons for the different results between the United States and Japan include (*a*) a higher percentage of office workers in females in the United States than in Japan, 31·9%, and 14·6%, respectively in 1975; (*b*) a higher divorce rate in the United States than in Japan, 4·1%, and 1·1% in 1975 respectively, indicating the need for consideration of smoking habits of ex-husbands in the United States; (*c*) smaller room size in Japan; (*d*) absence of the custom of asking, "Do you mind if I smoke?" in Japan; (*e*) other confounding variables.

In the survey conducted in Aichi Prefecture, one of our study areas, the number of passively smoked cigarettes and the length of exposure in the household was measured to estimate the quantity of sidestream smoke inhaled by the wives (table III). The number of cigarettes passively

TABLE I—*Mortality for lung cancer in non-smoking women by age, occupation,* and smoking habit of the husband*

Husband's age	Husband's occupation	Husband's smoking habit: Non-smoker — No of deaths	Non-smoker — Population	Ex-smoker or 1-19 cigarettes/day — No of deaths	Ex-smoker or 1-19 cigarettes/day — Population	≥20 cigarettes/day — No of deaths	≥20 cigarettes/day — Population	χ	Two-tailed p value
40–49	Total	5	6229	12	15034	15	10764	1·367	0·17163
	1	0	324	0	653	0	566		
	2	0	90	0	231	1	293	0·932	0·35134
	3	1	908	1	2247	3	1867	0·650	0·51569
	4	1	476	0	993	0	1044	−1·646	0·09976
	5	1	2502	6	5941	9	3636	2·297	0·02162
	6	0	46	0	185	0	108		
	7	0	177	1	486	1	426	0·541	0·58851
	8	1	1112	3	3431	1	2241	−0·545	0·59575
	9	0	162	1	345	0	243	−0·148	0·88234
	10	1	432	0	542	0	340	−1·217	0·22360
50–59	Total	6	7791	28	15642	21	9820	2·164	0·03046
	1	0	345	0	593	2	446	1·740	0·08186
	2	0	175	0	253	1	319	1·020	0·30773
	3	1	817	4	1764	1	1324	−0·436	0·66284
	4	1	653	2	1133	4	1092	0·955	0·33958
	5	2	3497	14	6812	7	3514	1·455	0·14567
	6	0	35	0	89	0	50		
	7	0	120	0	273	1	234	1·120	0·26271
	8	2	1375	6	3478	2	2155	−0·502	0·61567
	9	0	164	0	378	1	251	1·243	0·21387
	10	0	610	2	869	2	435	1·411	0·10718
60–69	Total	16	7120	42	12643	19	4651	1·790	0·07345
	1	0	227	0	327	1	179	1·434	0·15157
	2	1	91	0	143	0	124	−1·413	0·15766
	3	0	305	2	594	2	327	1·341	0·17992
	4	2	508	5	822	1	500	−0·459	0·64623
	5	11	4084	31	6845	8	2152	0·918	0·35862
	6	0	9	0	31	0	14		
	7	0	45	0	82	0	55		
	8	0	805	3	1784	3	736	1·872	0·06121
	9	1	121	0	208	0	92	−1·311	0·18986
	10	1	925	1	1607	4	472	2·391	0·01680
70 and over	Total	5	755	4	1065	1	226	−0·694	0·48768
	1	0	32	0	30	0	5		
	2	0	21	0	14	0	4		
	3	0	18	1	36	0	8	0·255	0·79872
	4	0	48	0	73	0	20		
	5	3	323	1	444	0	89	−1·494	0·13518
	6	0	1	0	1	0	0		
	7	0	1	0	5	0	1		
	8	0	87	2	119	1	36	1·371	0·17038
	9	0	11	0	19	0	2		
	10	2	213	0	322	0	61	−1·672	0·09452
Standardised risk ratio for all groups		1·000		1·435		1·846		2·942	0·00326

*Occupation: 1—Professional and technical workers; 2—managers and officials; 3—clerical and related workers; 4—sales workers; 5—farmers, lumbermen, and fishermen; 6—workers in mining and quarrying occupations; 7—workers in transport and communicating occupations; 8—craftsmen, production process workers, and labourers; 9—service workers; 10—not classifiable and not reported.

TABLE II—*Observed and expected deaths from lung cancer in non-smoking women with smoking husbands*

	Observed	Expected	Difference	Ratio	χ²	
Japan (Hirayama)	142	85·8	+56·2	+65·5%	36·81	Significant
USA (Garfinkel[1])	88	75·3	+12·7	+16·9%	2·14	Not significant
Greece (Trichopoulos *et al*[2])	29	12·1	+16·9	+139·7%	23·60	Significant
Total	259	173·2	+85·8	+49·5%	42·50	Significant

Source: https://www.industrydocuments.ucsf.edu/docs/lznx0000

2023381523